AD_____
(Leave blank)

Award Number: W81XWH-11-1-0024

TITLE:

Expression of Proteins Involved in Epithelial-Mesenchymal Transition as Predictors of Metastasis and Survival in Breast Cancer Patients

PRINCIPAL INVESTIGATOR: Michelle Roberts

CONTRACTING ORGANIZATION:

Health Research Inc.,Roswell Park Division

Buffalo NY 14263-0001

REPORT DATE: November 2013

TYPE OF REPORT: Revised Annual Summary Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
November 2013	Revised AnnualSummary	1Nov2012 to 31Oct2013

4. TITLE AND SUBTITLE
Expression of Proteins Involved in Epithelial-Mesenchymal Transition as Predictors of Metastasis and Survival in Breast Cancer Patients

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-1-0024

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Michelle Roberts, Wiam Bshara, Michael Higgins, Jo Freudenheim, Gokul Das, Austin Miller, and Christine Ambrosone

email: michelle.roberts@roswellpark.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Health Research Inc.,Roswell Park Division

Buffalo, NY 14263-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The purpose of this research is to investigate protein expression and promoter region DNA methylation status of six genes involved in epithelial-mesenchymal transition in relation to breast cancer lymph node metastasis, recurrence, and survival. Breast tumor tissue has been retrospectively identified from the Pathology Resource and immunohistochemical staining for all proteins has been completed. Scoring and analysis of IHC assays is ongoing. FFPE tissue cores for a subset of the IHC-stained tissues have been procured for isolation of DNA and methylation assay.

15. SUBJECT TERMS
Breast cancer; molecular epidemiology; epithelial-mesenchymal transition; metastasis; metastasis suppressor genes; immunohistochemistry; DNA methylation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	19	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Introduction: Breast cancer is incurable upon metastasis to distant organs, and metastasis to axillary lymph nodes is regarded as a critical prognostic factor for future recurrence and survival. Understanding the epidemiology and biology of metastasis could lead to better stratification of recurrence risk. We proposed to study genes related to epithelial-mesenchymal transition (EMT), invoking the hypotheses that EMT may explain the ability of tumor cells to form metastatic lesions and that these genes are regulated via DNA methylation. It is hypothesized that tumor cells co-opt the EMT program to transiently acquire properties generally reserved for mesenchymal cells, namely the ability to detach and migrate. The objectives of this project are to interrogate the protein expression and promoter methylation of six EMT-related genes: E-cadherin, N-cadherin, Vimentin, Twist1, RelB, and SATB1. Protein expression has been measured via immunohistochemistry (IHC) in breast tumor tissue and promoter methylation will be measured using DNA derived from these tumor samples. Protein expression and methylation status will be correlated with lymph node metastasis at diagnosis, time to metastatic recurrence, and disease-free survival. Effect modification by tumor grade, hormone receptor status, and HER2 status will also be investigated. This annual report describes the training and research accomplishments associated with the tasks outlined in the Statement of Work.

Training Plan

Tasks 1 and 5. All predoctoral program requirements have been completed. Dissertation defense planned for March 2014.

Task 2. Ongoing attendance in journal clubs for the Cancer Prevention research group and the Epigenetics research group at Roswell Park Cancer Institute (RPCI), works-in-progress meetings, weekly Institute-wide seminar series, and monthly Breast Disease Site Research Group meetings. Attendance at the 2013 American Association for Cancer Research (AACR) Annual Meeting.

Task 3. Continuation of research in molecular epidemiology focusing on molecular and genetic factors relating to lymph node metastasis, recurrences, and survival.

An analysis of the effect of polymorphisms in metastasis-related genes on the risk of lymph node positive breast cancer in African-American and European-American women is ongoing. We expect to submit a manuscript for publication within the next few months.

One paper was published in the journal Breast Cancer Research and Treatment in June 2013 (Appendix 1; Roberts M et al. Case-only analyses of the associations between polymorphisms in the metastasis modifying genes BRMS1 and SIPA1 and breast tumor characteristics, lymph node metastasis, and survival. Breast Cancer Research and Treatment 2013; 139:3:873-85.). In this paper, we examined the relationships between seven single nucleotide polymorphisms (SNPs) and lymph node status, tumor characteristics, overall survival, and recurrence-free survival in a cohort of 859 women diagnosed with invasive breast cancer, who were enrolled in

the Data Bank and BioRepository at Roswell Park Cancer Institute. We found that lymph node positive tumors were less likely among patients with the SIPA1 rs3741378 variant genotype, and more likely among patients heterozygous for the BRMS1 rs1052566 variant (Table 2). Having the variant genotype of SIPA1 rs7894763 was associated with an increased risk of high grade tumors (Table 3). Table 4 shows associations between the SNPs and tumor subtype. The variant genotype of BRMS1 rs3116068 was associated with an increased risk of having the luminal B or HER2-enriched tumor subtypes, while the BRMS1 rs1052566 variant was associated with a reduced risk of the luminal B tumor subtype. The variant genotypes of SIPA1 rs746429 and rs2306364 were associated with reduced risk of the triple negative subtype. We did not observe any significant associations with survival or recurrence (Table 5). Finally, to assess the effects of these SNPs together, we created a summary risk allele score (Table 6). We found that having 8 or more risk alleles was associated with significantly increased risk of lymph node positive tumor, and that overall, there was a dose-response relationship between the number of risk alleles and likelihood of node positivity (Ptrend = 0.002). There were no significant associations between the summary score and tumor grade or the survival outcomes, however.

Research Plan

Task 1. Interpretation of IHC assays is ongoing, using a combination of manual scoring and image analysis algorithms. This work is expected to be completed early in 2014, with data analysis expected to be completed during summer 2014.

Task 2. Conduct the laboratory assays necessary for Specific Aim 3, which proposes to evaluate gene promoter methylation of the E-cadherin, N-cadherin, Vimentin, Twist1, RelB, and SATB1 genes. DNA derived from the same FFPE tissues used in IHC was acquired from the Pathology Core Facility. We examined these samples for quality and quantity by several different methods, and found that the samples contained insufficient DNA for methylation analysis. We therefore requested new FFPE cores for DNA preparation. Cores for a subset of the patient population have been received and DNA preparation is planned for January 2014. Methylation analysis is expected to be completed during summer 2014.

Key Research Accomplishments:

None in this reporting period.

Reportable Outcomes:

One manuscript published in Breast Cancer Research and Treatment (Appendix 1; Roberts MR et al. Case-only analyses of the associations between polymorphisms in the metastasis modifying genes BRMS1 and SIPA1 and breast tumor characteristics, lymph node metastasis, and survival. Breast Cancer Research and Treatment 2013; 139:3:873-85.)

Conclusion: A one year no-cost extension was requested and approved, which will allow us to complete analysis of IHC data and complete the methylation assays. We expect this work to be completed during the summer months of 2014, with manuscript submission late in 2014.

Breast Cancer Res Treat (2013) 139:873–885
DOI 10.1007/s10549-013-2601-3

EPIDEMIOLOGY

Case-only analyses of the associations between polymorphisms in the metastasis-modifying genes *BRMS1* and *SIPA1* and breast tumor characteristics, lymph node metastasis, and survival

Michelle R. Roberts · Chi-Chen Hong · Stephen B. Edge · Song Yao · Wiam Bshara · Michael J. Higgins · Jo L. Freudenheim · Christine B. Ambrosone

Received: 15 March 2013 / Accepted: 6 June 2013 / Published online: 16 June 2013

Abstract Lymph node metastases and tumor characteristics predict breast cancer prognosis but correlate imperfectly with likelihood of metastatic relapse. Discovery of genetic polymorphisms affecting metastasis may improve identification of patients requiring aggressive adjuvant therapy to prevent recurrence. We investigated associations between several variants in the *BRMS1* and *SIPA1* metastasis-modifying genes and lymph node metastases, tumor subtype and grade, recurrence, disease-free survival, and overall survival. This cross-sectional and prospective prognostic analysis included 859 patients who received surgery for incident breast cancer at Roswell Park Cancer Institute, participated in the DataBank and BioRepository shared resource, and had DNA, clinical, and pathology data available for analysis. Genotyping for *BRMS1* (rs11537993, rs3116068, and rs1052566) and *SIPA1* (rs75894763, rs746429, rs3741378, and rs2306364) polymorphisms was performed using Sequenom® iPLEX Gold and Taqman® real-time PCR assays. Logistic and Cox proportional hazards regressions were used to estimate odds ratios (OR) and hazard ratios (HR), respectively. *BRMS1* rs1052566 heterozygous individuals were more likely to have node-positive tumors (OR = 1.58, 95 % CI 1.13–2.23), although there was no dose–response relationship, and those with at least one variant allele were less likely to have the luminal B subtype (AG + AA: OR = 0.59, 95 % CI 0.36–0.98). *BRMS1* rs3116068 was associated with increased likelihood of having the luminal B and the HER2-enriched tumor subtype ($P_{trend} = 0.03$). Two *SIPA1* SNPs, rs746429 and rs2306364, were associated with decreased risk of triple-negative tumors ($P_{trend} = 0.04$ and 0.07, respectively). Presence of 8 or more risk alleles was associated with an increased likelihood of having a node-positive tumor (OR = 2.14, 95 % CI 1.18–3.36, $P_{trend} = 0.002$). There were no significant associations with survival. Polymorphisms in metastasis-associated genes may be related to tumor characteristics and lymph node metastasis, but not survival. Future

M. R. Roberts (✉) · C.-C. Hong · S. Yao · C. B. Ambrosone
Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA
e-mail: michelle.roberts@roswellpark.org

C.-C. Hong
e-mail: chi-chen.hong@roswellpark.org

S. Yao
e-mail: song.yao@roswellpark.org

C. B. Ambrosone
e-mail: christine.ambrosone@roswellpark.org

M. R. Roberts · J. L. Freudenheim
Department of Social and Preventive Medicine, University at Buffalo, Buffalo, NY 14226, USA
e-mail: jfreuden@buffalo.edu

S. B. Edge
Department of Surgical Oncology, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA
e-mail: stephen.edge@roswellpark.org

W. Bshara
Department of Pathology, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA
e-mail: wiam.bshara@roswellpark.org

M. J. Higgins
Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA
e-mail: michael.higgins@roswellpark.org

evaluation of metastasis-modifying gene variants is necessary to better understand the biology of metastasis.

Keywords Breast cancer · Metastasis · Single nucleotide polymorphism · Recurrence · Survival

Abbreviations

BMI	Body mass index
BRMS1	Breast cancer metastasis suppressor 1
CI	Confidence interval
DBBR	DataBank and BioRepository
DCIS	Ductal carcinoma in situ
DFS	Disease-free survival
ER	Estrogen receptor
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
HRT	Hormone replacement therapy
NCCN	National Comprehensive Cancer Network
OR	Odds ratio
OS	Overall survival
PR	Progesterone receptor
RPCI	Roswell Park Cancer Institute
SIPA1	Signal-induced proliferation-associated 1
SNP	Single nucleotide polymorphism
TTR	Time to recurrence

Introduction

While early stage breast cancer has excellent prognosis, it is incurable once distant metastasis has occurred [1, 2]. Metastasis to regional lymph nodes is correlated with a higher risk of developing distant metastases [3, 4], as are tumor size and grade, estrogen and progesterone receptor status (ER and PR, respectively), and HER2 amplification. In general, larger tumors are correlated with increased likelihood of lymph node metastases at diagnosis and distant metastases [5, 6], while ER, PR, and HER2 status are markers of tumor aggressiveness and also determine suitability for targeted treatments [7].

Even with these known prognostic factors, however, the patients who will ultimately experience a recurrence are not clearly identified. Genetic variability may explain some of this heterogeneity in metastatic ability, particularly in genes affecting the metastatic cascade. Many metastasis-related genes have been identified, two of which are *BRMS1* (breast cancer metastasis suppressor 1) and *SIPA1* (signal-induced proliferation-associated 1).

BRMS1 can function as a metastasis suppressor gene [8–10] that affects apoptosis, colonization, cell adhesion, and invasive potential by mitigating the effects of anti-apoptotic gene products regulated by the NF-kB pathway [11, 12]. No studies examining single nucleotide polymorphisms (SNPs) in *BRMS1* have been published, although several expression studies have analyzed the relationship between *BRMS1* and breast tumor characteristics and prognosis [13–17]. *SIPA1* can affect metastatic efficiency by modifying cell adhesion [18] and expression of extracellular matrix genes [19] and has been shown to promote metastasis in vivo [20]. Several *SIPA1* SNP association studies have been published, with conflicting reports with respect to prognosis [21–23]. These data indicate that *BRMS1* and *SIPA1* abnormalities could affect tumor aggressiveness, metastasis, and the risk of recurrence in breast cancer patients.

Based on this previously published data, we selected several SNPs in *BRMS1* and *SIPA1* to investigate as potential candidate markers of tumor aggressiveness and recurrence. To investigate these relationships, we analyzed three SNPs in *BRMS1* [rs11537993 (Leu67Leu); rs3116068 (3′ UTR); and rs1052566 (Ala273Val)] and four SNPs in *SIPA1* [rs75894763 (Val621Val); rs746429 (Ala920Ala); rs3741378 (Ser182Phe); and rs2306364 (Ala342Ala)] with respect to lymph node metastasis, tumor grade and subtype, time to recurrence, disease-free survival, and overall survival in women diagnosed with primary, incident breast cancer.

Methods

Study population and outcomes

We identified 859 women diagnosed between October 2003 and May 2010 with stage I–III incident, primary, histologically confirmed breast cancer, who received surgery and treatment at Roswell Park Cancer Institute (RPCI), provided informed consent to RPCI's DataBank and BioRepository (DBBR), and had a DNA sample available. The DBBR, as previously described [24], is a comprehensive data and sample bank containing high-quality pre-treatment biospecimens and associated clinical and epidemiologic data. All patients diagnosed with cancer at RPCI are invited to participate. After consent and prior to treatment, including surgery, blood samples are collected, processed, and aliquoted for storage in liquid nitrogen. Epidemiologic data obtained via self-administered questionnaire were available for 688 of the participants in this analysis.

Outcomes were lymph node metastases, tumor subtype, tumor grade, time to recurrence, disease-free survival, and overall survival. Time to recurrence was defined as the time from diagnosis to date of first recurrence (local and regional recurrence and development of distant metastases) or last follow-up. Disease-free survival was defined as the

time from diagnosis to the date of first recurrence, death from any cause, or last follow-up. Overall survival was defined as the time from diagnosis to the date of death from any cause or last follow-up. Clinical data were obtained from RPCI clinical databases and supplemented with data abstracted from medical records and the RPCI Tumor Registry. Vital status and recurrence data were obtained from the RPCI Tumor Registry and the National Comprehensive Cancer Network (NCCN) Breast Cancer Outcomes Database. The RPCI Tumor Registry conducts yearly follow-up on patients who were last seen at RPCI 13 months prior and known to be alive. Vital status and recurrences are ascertained via RPCI medical record abstraction, Social Security Death Index and Legacy.com searches, and/or letters sent to the patient, the patient's physician, or a family member. NCCN-coordinated linkage with the National Death Index for patients defined as "lost to follow-up" was completed on 8 December 2011.

Fifteen participants missing HER2 status and 1 missing ER and PR status could not be classified by subtype. Additionally, 10 participants were missing lymph node status and 15 were missing tumor grade. Vital status was available for all participants. Follow-up ended in July 2012 and follow-up time ranged from 4 to 101 months. This study was approved by the RPCI Institutional Review Board.

Genotyping

The NCBI dbSNP resource was used to identify SNPs in the *BRMS1* and *SIPA1* genes [25]. We initially selected 13 SNPs in *BRMS1* [rs17850564 (Asp175Asp); rs11537993 (Leu67Leu); rs75053504 (A/G); rs3116068 (A/G); and rs1052566 (C/T)) and *SIPA1* (rs3741378 (Ser182Phe); rs76570058 (Pro1038Thr); rs75861149 (Gly368Gly); rs2306364 (Ala342Ala); rs75894763 (Val621Val); rs746429 (Ala920Ala); rs77600626 (Gly249Glu); and rs76089059 (Ala997Ala)] for genotyping, based on presence in protein coding, 3′ untranslated, or promoter regions, and heterozygosity of $\geq$0.10.

Genotyping of all 13 SNPs was conducted by the RPCI Genomics Facility using Sequenom MassARRAY® iPLEX Gold matrix-assisted laser desorption-ionization time-of-flight mass spectrometry assays. Genotyping of several SNPs (rs1052566, rs746429, rs3471378, rs2306364, rs77600626, and rs76089059) was unsuccessful using this platform, and an additional four SNPs (rs17850564, rs75053504, rs75861149, and rs76570058) were monomorphic and not analyzed further. Probes for Taqman® (Applied Biosciences) real-time PCR genotyping assays were available for four of the SNPs that failed Sequenom® genotyping (rs1052566, rs746429, rs3471378, and rs2306364). Therefore, we were ultimately able to obtain genotyping data for analysis of seven SNPs using either the Sequenom® (rs11537993, rs3116068, rs75894763) or Taqman® (rs1052566, rs746429, rs3471378, and rs2306364) platforms. Two SNPs in *SIPA1*, rs746429, and rs2306364 were in strong linkage disequilibrium ($r^2 = 0.809$). Duplicate samples were genotyped to assess intra- and inter-plate reliability. Genotyping call rates ranged from 96.2 to 99.8 %.

Cross-sectional analysis

Statistical analyses were performed using SAS version 9.3. Demographic variables and tumor characteristics were compared by lymph node status, tumor subtype, and tumor grade using Chi-squared and Fisher's exact tests as appropriate. Complete-case regression techniques were used to analyze the relationships between SNPs and lymph node metastasis, tumor grade, and tumor subtype. Potential covariates included age at diagnosis, tumor size, tumor grade, ER, PR, and HER2 status, lymph node metastasis, race, education, body mass index (BMI), age at menarche, menopausal status, age at menopause, parity, age at first birth, family history of breast cancer, history of benign breast disease, and hormone replacement therapy (HRT) use. Primary analyses incorporated data from the entire study population of 859 participants. Because we conducted complete-case analyses, participants missing epidemiologic questionnaire data dropped out of models including epidemiologic covariates. To minimize bias potentially introduced by these missing data, we initially included only tumor characteristics in adjusted models, as these data were available for the majority of our sample. Epidemiologic variables were included in separate models to assess the effect of their inclusion on odds ratio estimates. A variable was included in adjusted models if it was associated with the outcome and/or SNP(s), using Chi-squared or Fisher's exact tests of significance. Participants with and without questionnaire data had similar distributions of tumor characteristics and treatment modalities. Participants missing questionnaires were slightly younger and somewhat more likely to have node-positive tumors, but differences were not significant. For all outcomes, sensitivity analyses in which we excluded participants who self-identified as non-white (5.4 %) were performed.

Unconditional logistic regression was used to estimate odds ratios (OR) and 95 % confidence intervals (CI) for the associations between each of the seven SNPs and lymph node status and tumor grade. We first constructed age-adjusted models and then subsequently added ER and PR status, tumor size, and tumor grade. A third model additionally included HER2 status, race, education, HRT, and menopausal status. Finally, we restricted our analysis to include only stage 2 and 3 participants, as these patients are

eligible to have lymph node metastases (by definition, stage 1 is node negative).

Moderate- and low-grade tumors were combined, creating a dichotomous grade variable with categories of low/moderate grade (well-differentiated and moderately differentiated tumors) and high grade (poorly differentiated and undifferentiated tumors). Using a similar strategy as outlined above, we first adjusted for age and tumor size. In separate models, we added nodal status, ER, PR, and HER2 status, race, HRT, and menopausal status.

Generalized logit multinomial logistic regression was used to examine associations between each SNP and tumor subtype, using the luminal A subtype as the comparison group. Adjusted models included age and tumor size; age, tumor size, and nodal status; and age, tumor size, nodal status, HRT, race, and menopausal status.

For all analyses, P_{trend} was calculated by coding genotypes as 0, 1, or 2 for homozygous wild-type, heterozygous, or homozygous variant genotypes, respectively, and treating the SNP as a continuous variable in the regression models.

Survival analysis

Cox proportional hazards regression was used to estimate hazard ratios (HR) and 95 % confidence intervals (CI) for the relationships between the SNPs and overall survival, time to recurrence, and disease-free survival. Log-rank tests were used to identify predictors for inclusion in multivariate proportional hazards regression models. Variables tested as predictors were age at diagnosis, tumor size, tumor grade, ER, PR, and HER2 status, race, education, BMI, age at menarche, menopausal status, parity, family history of breast cancer, history of benign breast disease, hormone replacement therapy use, radiation treatment, chemotherapy, hormonal treatment (tamoxifen, etc.), and Charlson Comorbidity Score.

Significant variables were age, ER and PR status, tumor grade, tumor size, comorbidity score, radiation treatment, chemotherapy, hormone treatment, education, age at menarche, and menopausal status. Between 7 and 12 events, depending on the outcome, occurred among participants missing chemotherapy, hormone treatment, or comorbidity score. Adjustment for these variables would therefore result in loss of a large number of events. Similarly, adjustment for epidemiologic variables resulted in a large proportion of participants dropping out of our analyses due to missing data. To minimize bias due to dropout and maximize power, we initially limited model covariates to age, ER, PR, tumor grade, tumor size, and radiation treatment. We then tested the addition of other covariates (comorbidity score, chemotherapy, hormone therapy, education, age at menarche, and menopausal status) to assess their impact on the SNP–survival outcome associations. Finally, we conducted sensitivity analyses in models adjusted for age, ER, PR, radiation, tumor grade, and tumor size by testing the effect of including BMI and excluding non-white participants on hazard ratio estimates.

Risk allele score construction

We constructed a summary risk allele score using the log-additive model to estimate unadjusted per copy variant allele odds ratios and hazard ratios for lymph node status, tumor grade, time to recurrence, disease-free survival, and overall survival. We did not include tumor subtype in this analysis due to the complexity of creating a summary score for each subtype. Because the per copy variant allele effects were small in several instances, we considered odds ratios/hazard ratios that fell within the range 0.95–1.05 as being too close to null to assign a risk allele score. When this occurred, 0 risk alleles were assigned for all genotypes. If the per copy variant allele odds ratio/hazard ratio for a given SNP was greater than or equal to 1.06, genotypes for that SNP were assigned a score based on the following scheme: homozygous wild-type genotype = 0 risk alleles; heterozygous genotype = 1 risk allele; homozygous variant genotype = 2 risk alleles. If the odds ratio/hazard ratio was less than or equal to 0.94, the coding scheme was reversed: homozygous wild type = 2 risk alleles; heterozygous = 1 risk allele; homozygous variant = 0 risk alleles. The number of risk alleles for each of the seven SNPs was then added together for each participant to create the summary risk allele score. The directions of the log-additive odds and hazard ratios for each SNP were not generally similar across the lymph node status, tumor grade and survival analyses, leading to the assignment of different numbers of risk alleles for each outcome. Using the distributions of risk alleles for the lymph node status, tumor grade, and survival analyses, we categorized the number of risk alleles as 5 or less, 6, 7, and 8 or more to create a summary risk allele score. The category "5 or less" risk alleles served as the reference category. We then estimated the odds/hazard ratios for each level of the risk allele score, using logistic and proportional hazards regression as described above. Regression models included the same covariates as described previously for the lymph node status, tumor grade, and survival analyses.

Results

Participant characteristics are shown in Table 1. Younger participants were more likely to have node-positive, high-grade tumors of the luminal B, HER2-enriched, and triple-negative subtypes. Higher educational attainment was associated with decreased likelihood of node-positive

tumors. Ever users of HRT were more likely to have luminal B and HER2-enriched tumors. There were no other significant differences in demographic and reproductive variables with respect to tumor characteristics. Lymph node metastases were more commonly observed in conjunction with high-grade, larger size, ER-/PR-negative, and HER2-positive tumors.

SNP and lymph node status associations are presented in Table 2. *BRMS1* rs11537993 and *SIPA1* rs75894763, rs746429, and rs2306364 were not significantly associated with lymph node metastases. Age-adjusted odds ratios were similar to those obtained in multivariate models (data not shown). *BRMS1* rs1052566 heterozygotes were more likely to have node-positive tumors (OR = 1.58, 95 % CI 1.13–2.23), which remained significant after additional adjustment for race, HRT use, education, and menopausal status (OR = 1.70, 95 % CI 1.13–2.55) and when we limited the analysis to participants with stage 2 and 3 tumors (OR = 1.85, 95 % CI 1.08–3.18). However, this relationship was not observed among homozygous individuals. Although only marginally significant, participants with at least one copy of the variant *SIPA1* rs3741378 allele were less likely to have node-positive tumors (OR = 0.70, 95 % CI 0.48–1.02). While the direction and magnitude persisted following adjustment for additional covariates and limitation to stage 2 and 3 tumors, this association became nonsignificant. Similarly, *BRMS1* rs3116068 approached statistical significance only when race, HRT use, education, and menopausal status were added as covariates (OR = 0.67, 95 % CI 0.45–1.01).

SNP associations with tumor grade are presented in Table 3. *SIPA1* rs75894763 heterozygous participants were more likely to have high-grade tumors (OR = 2.62, 95 % CI 1.06–6.48), but only after further adjustment for race, HRT use, and menopausal status.

Associations between tumor subtype and genotype are shown in Table 4. Results of age- and tumor size-adjusted analyses were similar to the findings presented in Table 4 and are not shown. The *BRMS1* rs3116068 homozygous variant genotype was associated with increased likelihood of luminal B tumors (OR = 2.50, 95 % CI 1.10–5.66), although there was a nonsignificant inverse relationship among heterozygotes. Those who were heterozygous or homozygous were also more likely to have tumors of the HER2-enriched subtype (OR = 2.45, 95 % CI 1.18–5.06, $P_{\text{trend}} = 0.03$). These relationships remained significant after additional adjustment for race, HRT use, and menopausal status (data not shown). Patients homozygous or heterozygous for *BRMS1* rs1052566 were less likely to have luminal B tumors (OR = 0.59, 95 % CI 0.36–0.98,

Table 1 Participant characteristics by lymph node status, tumor subtype, and tumor grade

Characteristic[a], *n* (%)	Lymph node status[b]		Tumor subtype[b]				Tumor grade[b]		
	Positive (*N* = 246)	Negative (*N* = 603)	Luminal A (*N* = 596)	Luminal B (*N* = 75)	HER2 (+) (*N* = 34)	Triple (−) (*N* = 138)	Low (*N* = 225)	Moderate (*N* = 371)	High (*N* = 248)
Age at diagnosis									
≤50	97 (39.4)	177 (29.4)*	174 (29.2)	29 (38.7)	15 (44.1)	51 (37.0)	62 (27.6)	115 (31.0)	93 (37.5)
51–65	89 (36.2)	253 (42.0)	246 (41.3)	28 (37.3)	15 (44.1)	50 (36.2)	95 (42.2)	141 (38.0)	99 (39.9)
≥66	60 (24.4)	173 (28.7)	176 (29.5)	18 (24.0)	4 (11.8)	37 (26.8)	68 (30.2)	115 (31.0)	56 (22.6)
Race									
White	171 (90.5)	463 (94.5)	450 (94.5)	60 (92.3)	19 (86.4)	103 (91.2)	172 (96.1)	277 (93.9)	184 (90.6)
Non-white	18 (9.5)	27 (5.5)	26 (5.5)	5 (7.7)	3 (13.6)	10 (8.8)	7 (3.9)	18 (6.1)	19 (9.4)
Missing	57	113	120	10	12	25	46	76	45
Education									
High school or less	79 (41.8)	166 (34.5)	180 (38.4)	19 (29.7)	8 (36.4)	39 (34.8)	63 (36.0)	109 (37.3)	76 (37.8)
At least some college	110 (58.2)	315 (65.5)	289 (61.6)	45 (70.3)	14 (63.6)	73 (65.2)	112 (64.0)	183 (62.7)	125 (62.2)
Missing	57	122	127	11	12	26	50	79	47
Menopausal status									
Pre	72 (38.7)	157 (32.2)	150 (31.8)	29 (44.6)	7 (31.8)	39 (34.8)	58 (32.8)	92 (31.4)	77 (38.1)
Post	114 (61.3)	331 (67.8)	322 (68.2)	36 (55.4)	15 (68.2)	73 (65.2)	119 (67.2)	201 (68.6)	125 (61.9)
Missing	60	115	124	10	12	26	48	78	46
HRT use[c]									
Never	68 (60.7)	170 (53.1)	183 (58.8)	15 (42.9)	4 (26.7)	39 (54.2)*	66 (57.4)	113 (58.2)	62 (50.4)
Ever	44 (39.3)	150 (46.9)	128 (41.2)	20 (57.1)	11 (73.3)	33 (45.8)	49 (42.6)	81 (41.8)	61 (49.6)
Missing	2	11	11	1	0	1	4	7	2
Parity									
Nulliparous	28 (15.1)	94 (19.2)	84 (17.8)	17 (26.2)	3 (13.6)	14 (12.5)	34 (19.0)	51 (17.4)	35 (17.4)

Table 1 continued

Characteristic[a], *n* (%)	Lymph node status[b]		Tumor subtype[b]				Tumor grade[b]		
	Positive ($N = 246$)	Negative ($N = 603$)	Luminal A ($N = 596$)	Luminal B ($N = 75$)	HER2 (+) ($N = 34$)	Triple (−) ($N = 138$)	Low ($N = 225$)	Moderate ($N = 371$)	High ($N = 248$)
Parous	158 (84.9)	395 (80.8)	389 (82.2)	48 (73.8)	19 (86.4)	98 (87.5)	145 (81.0)	242 (82.6)	166 (82.6)
Missing	60	114	123	10	12	26	46	78	47
ER status									
Positive	185 (75.2)	479 (80.0)	590 (99.0)	73 (97.3)	0	0	220 (98.2)	332 (89.5)	113 (45.7)*
Negative	61 (24.8)	120 (20.0)	6 (1.0)	2 (2.7)	34	138	4 (1.8)	39 (10.5)	134 (54.3)
Missing	0	4	0	0	0	0	1	0	1
PR status									
Positive	154 (62.6)	420 (70.1)*	512 (85.9)	63 (84.0)	0	0	200 (89.3)	281 (75.7)	91 (36.8)*
Negative	92 (37.4)	179 (29.9)	84 (14.1)	12 (16.0)	34	138	24 (10.7)	90 (24.3)	156 (63.2)
Missing	0	4	0	0	0	0	1	0	1
HER2 status									
Positive	44 (17.9)	65 (11.0)*	0	75	34	0	4 (1.8)	46 (12.6)	55 (22.4)*
Negative	202 (82.1)	524 (89.0)	596	0	0	138	217 (98.2)	320 (87.4)	190 (77.6)
Missing	0	14 (2.3)	0	0	0	0	4	5	3
Lymph node status									
Positive	246	–	163 (27.7)	23 (30.7)	21 (61.8)	39 (28.5)*	47 (21.4)	106 (28.8)	90 (36.6)*
Negative	–	603	425 (72.3)	52 (69.3)	13 (38.2)	98 (71.5)	173 (78.6)	262 (71.2)	156 (63.4)
Missing	–	–	8	0	0	1	5	3	2
Tumor subtype									
Luminal A	163 (66.3)	425 (72.3)*	596	–	–	–	213 (96.8)	293 (80.1)	84 (34.3)*
Luminal B	23 (9.4)	52 (8.8)	–	75	–	–	4 (1.8)	37 (10.1)	32 (13.1)
HER2 (+)	21 (8.5)	13 (2.2)	–	–	34	–	0	9 (2.5)	23 (9.4)
Triple negative	39 (15.9)	98 (16.7)	–	–	–	138	3 (1.4)	27 (7.4)	106 (43.3)
Missing	0	15	–	–	–	–	5	5	3
Tumor grade									
Low	47 (19.3)	173 (29.3)*	213 (36.1)	4 (5.5)	0	3 (2.2)*	225	–	–
Moderate	106 (43.6)	262 (44.3)	293 (49.7)	37 (50.7)	9 (28.1)	27 (19.9)	–	371	–
High	90 (37.1)	156 (26.4)	84 (14.2)	32 (43.8)	23 (71.9)	106 (77.9)	–	–	248
Missing	3	12	6	2	2	2	–	–	–
Tumor size									
Tmi, T1A (≤5 mm)	14 (5.7)	89 (14.8)*	67 (11.2)	6 (8.0)	5 (14.7)	14 (10.1)*	37 (16.4)	40 (10.8)	18 (7.3)*
T1B, T1C (>5–20 mm)	116 (47.2)	413 (68.5)	398 (66.8)	45 (60.0)	14 (41.2)	78 (56.5)	159 (70.7)	239 (64.4)	136 (54.8)
T2 (>20–50 mm)	96 (39.0)	96 (15.9)	115 (19.3)	23 (30.7)	12 (35.3)	41 (29.7)	24 (10.7)	82 (22.1)	84 (33.9)
T3 (>50 mm)	20 (8.1)	5 (0.8)	16 (2.7)	1 (1.3)	3 (8.8)	5 (3.6)	5 (2.2)	10 (2.7)	10 (4.0)
Stage									
1	2 (0.9)	498 (83.4)	366 (62.8)	42 (56.0)	9 (26.5)	71 (51.8)*	156 (72.2)	225 (61.3)	111 (45.1)*
2	164 (66.9)	99 (16.6)	170 (29.2)	27 (36.0)	15 (44.1)	50 (36.5)	48 (22.2)	115 (31.3)	96 (39.0)
3	79 (32.2)	0	47 (8.0)	6 (8.0)	10 (29.4)	16 (11.7)	12 (5.6)	27 (7.4)	39 (15.9)
Missing	1	6	13	0	0	1	9	4	2

HRT hormone replacement therapy, *ER* estrogen receptor, *PR* progesterone receptor

* Significant at $\alpha = 0.05$; *p* values were obtained from Chi-squared and Fisher's exact tests as appropriate (missing categories are excluded from *p* value calculations)

[a] Race, education, menopausal status, HRT use, and parity were available for 688 participants; age at diagnosis, ER, PR, HER2 status, tumor grade, and tumor size were available for 859 participants

[b] Lymph node status, subtype, and grade were missing for $n = 10$, $n = 16$, and $n = 15$, respectively

[c] Excludes premenopausal women

Table 2 Associations of *BRMS1* and *SIPA1* SNPs with presence of lymph node metastases at diagnosis

Genotype			Odds ratio for likelihood of node-positive tumor at diagnosis				
			Node positive, *n*	Node negative, *n*	Adjusted[a] OR (95 % CI)	Adjusted[b] OR (95 % CI)	Adjusted[c] OR (95 % CI)
BRMS1	rs11537993	AA	125	293	1.00	1.00	1.00
	(Leu67Leu)	AG	98	243	0.93 (0.66–1.30)	1.02 (0.69–1.53)	0.98 (0.60–1.63)
		GG	20	50	0.90 (0.49–1.63)	1.01 (0.53–1.92)	1.14 (0.44–2.93)
		AG + GG			0.92 (0.67–1.27)	1.02 (0.70–1.49)	1.01 (0.62–1.63)
		P_{trend}			0.61	0.95	0.89
	rs3116068	GG	165	365	1.00	1.00	1.00
	(3′ UTR)	AG	66	183	0.80 (0.56–1.14)	0.66 (0.43–1.02)	0.69 (0.41–1.18)
		AA	10	34	0.66 (0.30–1.43)	0.73 (0.31–1.72)	0.59 (0.18–1.89)
		AG + AA			0.78 (0.55–1.09)	0.67 (0.45–1.01)	0.68 (0.41–1.13)
		P_{trend}			0.13	0.09	0.13
	rs1052566	GG	102	285	1.00	1.00	1.00
	(Ala273Val)	AG	115	220	1.58 (1.13–2.23)	1.70 (1.13–2.55)	1.85 (1.08–3.18)
		AA	21	61	0.92 (0.51–1.65)	1.06 (0.54–2.08)	0.72 (0.33–1.55)
		AG + AA			1.43 (1.03–1.99)	1.56 (1.06–2.30)	1.48 (0.91–2.42)
		P_{trend}			0.25	0.17	0.67
SIPA1	rs75894763	GG	237	558	1.00	1.00	1.00
	(Val621Val)	AG	5	29	0.39 (0.14–1.08)	0.51 (0.16–1.60)	0.30 (0.09–1.03)
	rs746429	GG	110	267	1.00	1.00	1.00
	(Ala920Ala)	GA	102	239	0.99 (0.70–1.40)	1.06 (0.71–1.59)	0.86 (0.52–1.43)
		AA	29	66	1.09 (0.65–1.83)	1.16 (0.64–2.12)	1.31 (0.55–3.14)
		GA + AA			1.01 (0.73–1.40)	1.08 (0.74–1.58)	0.93 (0.57–1.51)
		P_{trend}			0.82	0.61	0.88
	rs3741378	CC	187	413	1.00	1.00	1.00
	(Ser182Phe)	TC	49	148	0.71 (0.48–1.04)	0.75 (0.47–1.19)	0.73 (0.41–1.29)
		TT	4	13	0.62 (0.18–2.10)	0.61 (0.14–2.67)	0.42 (0.09–2.05)
		TC + TT			0.70 (0.48–1.02)	0.74 (0.47–1.16)	0.70 (0.40–1.21)
		P_{trend}			0.07	0.18	0.15
	rs2306364	GG	186	425	1.00	1.00	1.00
	(Ala342Ala)	AG	13	42	0.71 (0.35–1.41)	0.70 (0.34–1.48)	0.89 (0.30–2.63)
		AA	42	99	0.98 (0.64–1.50)	0.92 (0.56–1.53)	1.14 (0.58–2.24)
		AG + AA			0.90 (0.62–1.32)	0.85 (0.55–1.32)	1.07 (0.59–1.95)
		P_{trend}			0.78	0.62	0.75

Unconditional logistic regression was used to estimate odds ratios (OR) and 95 % confidence intervals (CI) for risk of node-positive tumors

[a] Adjusted for age at diagnosis, ER, PR, tumor size, and tumor grade

[b] Adjusted for age at diagnosis, ER, PR, HER2, tumor size, tumor grade, race, HRT, education, and menopausal status

[c] Limited to women with stage 2 and 3 breast cancer, adjusted for age at diagnosis, ER/PR status, and tumor grade. Tumor size was not included as a covariate due to sparse data

$P_{trend} = 0.05$). These relationships remained marginally significant after further adjustment for race, HRT use, and menopausal status (OR = 0.59, 95 % CI 0.34–1.02). A nonsignificant decrease in the likelihood of the HER2-enriched subtype was observed among patients heterozygous or homozygous for this variant, which became statistically significant when race, HRT use, and menopausal status were included in the model (OR = 0.32, 95 % CI 0.12–0.90). Participants homozygous for *SIPA1* rs746429 were less likely to have triple-negative tumors (OR = 0.48, 95 % CI 0.24–0.97, $P_{trend} = 0.04$). Similarly, participants either homozygous or heterozygous for *SIPA1* rs2306364 were less likely to have triple-negative tumors, although this relationship was only borderline significant (OR = 0.62, 95 % CI 0.38–1.01).

Associations with survival outcomes are shown in Table 5. The median follow-up time was 45 months (range 4–101 months), during which 58 recurrences and 70 deaths

Table 3 Associations of *BRMS1* and *SIPA1* SNPs with tumor grade

Genotype			Odds ratio for likelihood of high-grade tumor at diagnosis			
			High grade, *n*	Low/moderate grade, *n*	Adjusted[a] OR (95 % CI)	Adjusted[b] OR (95 % CI)
BRMS1	rs11537993	AA	128	286	1.00	1.00
	(Leu67Leu)	AG	92	243	0.86 (0.58–1.27)	0.97 (0.62–1.53)
		GG	23	47	1.23 (0.65–2.35)	1.36 (0.68–2.70)
		AG + GG			0.92 (0.64–1.33)	1.04 (0.68–1.60)
		P_{trend}			0.97	0.55
	rs3116068	GG	156	369	1.00	1.00
	(3′ UTR)	AG	70	175	0.83 (0.55–1.26)	0.98 (0.61–1.56)
		AA	11	32	0.82 (0.36–1.89)	0.67 (0.26–1.75)
		AG + AA			0.83 (0.56–1.23)	0.92 (0.59–1.44)
		P_{trend}			0.38	0.54
	rs1052566	GG	113	267	1.00	1.00
	(Ala273Val)	AG	92	240	0.86 (0.58–1.28)	0.97 (0.61–1.53)
		AA	27	55	1.39 (0.76–2.57)	1.63 (0.82–3.25)
		AG + AA			0.95 (0.66–1.38)	1.08 (0.70–1.66)
		P_{trend}			0.67	0.33
SIPA1	rs75894763	GG	233	553	1.00	1.00
	(Val621Val)	AG	10	23	1.97 (0.87–4.47)	2.62 (1.06–6.48)
	rs746429	GG	114	259	1.00	1.00
	(Ala920Ala)	GA	104	232	1.20 (0.81–1.78)	1.28 (0.82–2.00)
		AA	21	73	0.82 (0.44–1.54)	0.83 (0.41–1.70)
		GA + AA			1.11 (0.77–1.61)	1.17 (0.77–1.79)
		P_{trend}			0.98	0.91
	rs3741378	CC	166	429	1.00	1.00
	(Ser182Phe)	TC	65	127	1.44 (0.94–2.20)	1.46 (0.89–2.39)
		TT	6	11	0.99 (0.27–3.60)	1.43 (0.36–5.71)
		TC + TT			1.40 (0.92–2.11)	1.46 (0.90–2.35)
		P_{trend}			0.18	0.14
	rs2306364	GG	187	419	1.00	1.00
	(Ala342Ala)	AG	13	40	0.84 (0.40–1.77)	0.72 (0.32–1.61)
		AA	36	103	0.93 (0.57–1.53)	0.93 (0.53–1.63)
		AG + AA			0.90 (0.58–1.40)	0.86 (0.52–1.40)
		P_{trend}			0.71	0.67

Unconditional logistic regression was used to estimate odds ratios (OR) and 95 % confidence intervals (CI) for risk of high-grade tumors

Low grade = well differentiated, *Moderate grade* = moderately differentiated, *High grade* = poorly differentiated and undifferentiated

[a] Adjusted for age at diagnosis, tumor size, lymph node status, ER, PR, and HER2

[b] Adjusted for age at diagnosis, tumor size, lymph node status, ER, PR, HER2, race, HRT, and menopausal status

from all causes occurred. When adjusted for age only, the heterozygous genotype of *BRMS1* rs3116068 was associated with poorer overall survival (HR = 1.65, 95 % CI 1.02–2.68), but this association became nonsignificant when additional covariates were included. We did not observe any other significant associations. Results were unchanged when additional covariates (chemotherapy, hormone therapy, education, age at menarche, and menopausal status) were included (data not shown).

One of our hypotheses is that these SNPs would be related to lymph node status. Because nodal status is also related to survival outcomes, it can be hypothesized that this variable is part of the causal pathway, and therefore adjustment for nodal status could mask true SNP–survival associations. To test whether these SNPs could affect vital status or recurrence through a pathway independent of lymph node status, we included nodal status in a model containing age, ER, PR, radiation, tumor grade, and tumor

Table 4 Associations of *BRMS1* and *SIPA1* SNPs with tumor subtype

Genotype			Odds ratio for likelihood of luminal B, HER2 (+), and triple (−) subtype, compared to luminal A						
			Luminal A, *n*	Luminal B, *n*	Adjusted[b] OR (95 % CI)	HER2 (+), n	Adjusted[b] OR (95 % CI)	Triple (−), n	Adjusted[b] OR (95 % CI)
BRMS1	rs11537993	AA	292	33	1.00	18	1.00	77	1.00
	(Leu67Leu)	AG	246	31	1.14 (0.68–1.93)	15	1.13 (0.55–2.34)	48	0.74 (0.50–1.11)
		GG	49	10	1.85 (0.85–4.01)	1	0.36 (0.05–2.82)	11	0.74 (0.50–1.11)
		AG + GG			1.26 (0.77–2.06)		0.99 (0.49–2.03)		0.76 (0.52–1.12)
		P_{trend}			0.17		0.64		0.26
	rs3116068	GG	379	50	1.00	15	1.00	90	1.00
	(3′ UTR)	AG	174	16	0.71 (0.39–1.28)	16	2.49 (1.18–5.27)	41	1.00 (0.66–1.51)
		AA	28	9	2.50 (1.10–5.66)	2	2.11 (0.45–10.0)	5	0.76 (0.28–2.04)
		AG + AA			0.95 (0.57–1.59)		2.45 (1.18–5.06)		0.97 (0.65–1.44)
		P_{trend}			0.39		0.03		0.74
	rs1052566	GG	268	44	1.00	18	1.00	58	1.00
	(Ala273Val)	AG	240	24	0.61 (0.36–1.03)	13	0.70 (0.33–1.49)	57	1.12 (0.74–1.68)
		AA	61	6	0.55 (0.22–1.36)	1	0.21 (0.03–1.66)	15	1.07 (0.57–2.03)
		AG + AA			0.59 (0.36–0.98)		0.60 (0.29–1.25)		1.11 (0.75–1.63)
		P_{trend}			0.05		0.09		0.69
SIPA1	rs75894763	GG	559	69	1.00	34	1.00	135	1.00
	(Val621Val)	AG	27	5	1.47 (0.54–3.99)	0	NA	2	0.29 (0.07–1.26)
	rs746429	GG	255	35	1.00	19	1.00	69	1.00
	(Ala920Ala)	GA	245	29	0.85 (0.50–1.44)	11	0.56 (0.26–1.22)	55	0.81 (0.54–1.20)
		AA	75	9	0.86 (0.39–1.87)	3	0.51 (0.14–1.80)	10	0.48 (0.24–0.97)
		GA + AA			0.85 (0.52–1.39)		0.55 (0.27–1.13)		0.73 (0.50–1.07)
		P_{trend}			0.57		0.13		0.04
	rs3741378	CC	429	53	1.00	27	1.00	93	1.00
	(Ser182Phe)	TC	138	17	0.99 (0.55–1.78)	4	0.46 (0.16–1.37)	37	1.21 (0.79–1.87)
		TT	9	3	2.91 (0.75–11.2)	2	5.77 (1.11–29.8)	3	1.57 (0.41–5.98)
		TC + TT			1.11 (0.64–1.92)		0.68 (0.27–1.71)		1.24 (0.81–1.88)
		P_{trend}			0.45		0.90		0.30
	rs2306364	GG	421	58	1.00	27	1.00	107	1.00
	(Ala342Ala)	AG	42	6	1.02 (0.41–2.51)	1	0.38 (0.05–2.94)	6	0.56 (0.23–1.36)
		AA	109	9	0.59 (0.28–1.23)	5	0.74 (0.27–2.00)	18	0.64 (0.37–1.11)
		AG + AA			0.71 (0.39–1.29)		0.64 (0.26–1.60)		0.62 (0.38–1.01)
		P_{trend}			0.18		0.44		0.07

Multinomial logistic regression (generalized logit model) was used to estimate odds ratios (OR) and 95 % confidence intervals (CI) for risk of luminal B, HER2 (+), or triple-negative tumor subtype, using luminal A as the comparison group

Luminal A = ER and/or PR positive, HER2 negative; *Luminal B* = ER and/or PR positive, HER2 positive; *HER2-enriched subtype* = (HER2 (+)) ER and PR negative, HER2 positive; *Triple negative* = (Triple (−)) ER, PR, and HER2 negative

[a] Adjusted for age at diagnosis, tumor size, and lymph node status

size as covariates. There was no change in the hazard ratio estimates for any of the three outcomes (data not shown).

Results of the summary risk allele score analysis are shown in Table 6. Having eight or more risk alleles was associated with significantly greater likelihood of having a node-positive tumor (OR = 2.14 95 % CI 1.18–3.86). There was evidence of a dose–response pattern, with increasing numbers of risk alleles associated with increased likelihood of node positivity ($P_{trend} = 0.002$). There were no significant associations with tumor grade or the three survival outcomes.

In sensitivity analyses, non-white participants were excluded for all study outcomes, which did not alter our findings (data not shown).

Table 5 *BRMS1* and *SIPA1* SNP associations with time to recurrence, disease-free survival, and overall survival

Genotype			No. events, TTR	TTR, adjusted[a] HR (95 % CI)	No. events, DFS	DFS, adjusted[a] HR (95 % CI)	No. events, OS	OS, adjusted[a] HR (95 % CI)
BRMS1	rs11537993	AA	27	1.00	39	1.00	28	1.00
	(Leu67Leu)	AG	21	1.05 (0.59–1.89)	35	1.12 (0.70–1.79)	28	1.10 (0.64–1.89)
		GG	5	1.24 (0.47–3.30)	7	1.18 (0.52–2.69)	6	1.30 (0.53–3.20)
		AG + GG		1.08 (0.62–1.89)		1.13 (0.72–1.77)		1.13 (0.67–1.89)
		P_{trend}		0.69		0.59		0.56
	rs3116068	GG	35	1.00	48	1.00	33	1.00
	(3′ UTR)	AG	15	0.77 (0.41–1.43)	29	1.04 (0.65–1.67)	26	1.44 (0.85–2.45)
		AA	3	0.90 (0.26–3.13)	3	0.54 (0.16–1.82)	1	0.33 (0.04–2.49)
		AG + AA		0.78 (0.43–1.42)		0.97 (0.61–1.55)		1.31 (0.77–2.22)
		P_{trend}		0.51		0.60		0.76
	rs1052566	GG	26	1.00	41	1.00	33	1.00
	(Ala273Val)	AG	22	1.05 (0.59–1.87)	33	0.99 (0.62–1.57)	23	0.86 (0.50–1.48)
		AA	4	1.01 (0.35–2.96)	5	0.97 (0.38–2.48)	3	0.81 (0.24–2.67)
		AG + AA		1.04 (0.60–1.81)		0.98 (0.63–1.54)		0.85 (0.51–1.44)
		P_{trend}		0.92		0.93		0.55
SIPA1	rs75894763	GG	51	1.00	78	1.00	60	1.00
	(Val621Val)	AG	2	1.61 (0.38–6.89)	3	1.57 (0.48–5.11)	2	1.35 (0.32–5.75)
	rs746429	GG	24	1.00	37	1.00	28	1.00
	(Ala920Ala)	GA	25	1.20 (0.67–2.14)	38	1.15 (0.72–1.83)	30	1.24 (0.73–2.10)
		AA	5	0.94 (0.36–2.50)	7	0.86 (0.38–1.94)	4	0.69 (0.24–1.99)
		GA + AA		1.14 (0.66–1.99)		1.09 (0.70–1.70)		1.13 (0.67–1.89)
		P_{trend}		0.84		0.98		0.94
	rs3741378	CC	37	1.00	58	1.00	45	1.00
	(Ser182Phe)	TC	17	1.28 (0.71–2.30)	24	1.09 (0.67–1.77)	16	0.86 (0.48–1.55)
		TT	0	NA	0	NA	1	1.16 (0.16-8.63)
		TC + TT		1.20 (0.67–2.16)		1.03 (0.63–1.67)		0.87 (0.49–1.55)
		P_{trend}		NA		NA		0.70
	rs2306364	GG	41	1.00	61	1.00	48	1.00
	(Ala342Ala)	AG	5	1.16 (0.45–3.01)	7	1.20 (0.54–2.71)	4	0.92 (0.32–2.64)
		AA	8	0.82 (0.38–1.75)	13	0.91 (0.50–1.66)	9	0.80 (0.39–1.65)
		AG + AA		0.92 (0.49–1.73)		0.99 (0.59–1.66)		0.83 (0.45–1.56)
		P_{trend}		0.67		0.84		0.54

Patients whose recurrence status indicated that they were never disease free were excluded from time to recurrence and disease-free survival analyses, but were included in overall survival analyses

Hazard ratios (HR) and 95 % confidence intervals (CI) were estimated using Cox proportional hazards regression

TTR = Time to recurrence (time from diagnosis to date of first recurrence or date of last follow-up), *DFS* = disease-free survival (time from diagnosis to date of first recurrence, death, or last follow-up), *OS* = overall survival (time from diagnosis to date of death or date of last follow-up)

[a] Adjusted for age at diagnosis, ER, PR, tumor size, tumor grade, radiation treatment, and Charlson Comorbidity Score

Discussion

Our data suggest that 2 SNPs in the *BRMS1* gene, rs1052566 and rs3116068, may be associated with lymph node status and tumor subtype, and that SNPs in the *SIPA1* gene may be associated with tumor grade and subtype. We also found that a summary score, composed of the number of "at risk" alleles for each of the seven *BRMS1* and *SIPA1* SNPs analyzed, was significantly associated with lymph node status.

To our knowledge, associations between SNPs in *BRMS1* and breast tumor characteristics and prognosis have not been previously evaluated. *BRMS1* has multiple functions, including transcriptional regulation via NF-κB signaling pathways [26, 27], chromatin modification [26], interactions with histone deacetylase complexes [27], and

Table 6 Risk allele score associations with lymph node status, tumor grade, time to recurrence, disease-free survival, and overall survival

Number of alleles	Lymph node status			Tumor grade			Time to recurrence		Disease-free survival		Overall survival	
	Positive (*n*)	Negative (*n*)	OR[a] (95 % CI)	High (*n*)	Low/Mod (*n*)	OR[b] (95 % CI)	N*	HR[c] (95 % CI)	N*	HR[c] (95 % CI)	N*	HR[c] (95 % CI)
5 or less	18	73	1.00	153	362	1.00	19	1.00	15	1.00	34	1.00
6	50	148	1.31 (0.69–2.50)	53	150	0.85 (0.54–1.32)	21	1.83 (0.96–3.48)	32	1.84 (0.98–3.46)	20	1.19 (0.67–2.10)
7	44	117	1.38 (0.71–2.67)	33	61	1.19 (0.67–2.10)	12	1.22 (0.59–2.54)	15	0.86 (0.41–1.82)	8	1.74 (0.79–3.83)
8 or more	131	251	2.14 (1.18–3.86)	4	6	1.28 (0.28–5.91)	2	0.50 (0.11–2.17)	20	1.13 (0.57–2.24)	0	NA
P_{trend}			0.002			0.78		0.82		0.60		0.48

N is the number of participants in each risk allele score category

N* is the number of events in each risk allele score category

[a] Adjusted for age at diagnosis, ER, PR, tumor size, and tumor grade

[b] Adjusted for age at diagnosis, ER, PR, HER2, tumor size, and lymph node status

[c] Adjusted for age at diagnosis, ER, PR, tumor size, tumor grade, radiation treatment, and Charlson Comorbidity Score

transcriptional repression of anti-apoptotic genes [12]. Previous studies have correlated decreased *BRMS1* gene expression with breast tumor aggressiveness. Reduced mRNA and protein expression in breast tumors has been associated with PR-negative, HER2-positive tumors, as well as younger age at diagnosis [13, 14], but not with lymph node metastases [14, 15]. One study found that *BRMS1* mRNA was reduced in brain metastases of breast cancer patients [16]. In survival analyses, both increased and decreased *BRMS1* gene expression have been correlated with reduced survival [14, 17], while loss of *BRMS1* protein expression has been associated with reduced survival only in patients with ER-negative, HER2-positive tumors [13].

We observed several relationships between *SIPA1* SNPs rs746429, rs3741378, and rs2306364 and tumor subtype, although there are too few participants with variant alleles to draw strong conclusions. *SIPA1* encodes a GTPase-activating protein and affects expression of extracellular matrix genes [19]. *SIPA1* SNPs rs931127, rs3741378, and rs746429 have not been shown to be associated with overall survival [22], similar to our findings. *SIPA1* rs3741378 has been correlated with increased risk of ER/PR-negative tumors, while rs746429 has been correlated with increased risk of node-positive breast tumors [21]. In our study, rs3741378 was associated with an increased risk of HER2-enriched tumors, which are ER/PR negative.

Our study used data and samples collected under the DBBR's standardized protocol and had relatively large sample size, but our analyses were limited by the small number of outcomes and participants with the variant genotypes. We also examined only a few SNPs in each gene. Our ability to adjust for socioeconomic and reproductive covariates was limited by missing epidemiologic questionnaire data, although these factors are unlikely to be strong confounders as SNPs are generally unlikely to be associated with them. Distributions of tumor characteristics and treatment were similar between participants with and without questionnaire data.

We compared the characteristics of participants with complete data to the characteristics of those with incomplete data to assess the possibility of bias. In general, participants with incomplete data were more likely to be postmenopausal, non-white, to have smaller tumors, and to have received hormonal treatment, but were less likely to have a family history of breast cancer and to have received radiation and chemotherapy. Age, ER, PR, HER2, tumor grade, comorbidity score, parity, age at menarche, and history of benign breast disease were generally similar. This indicates that the participants who dropped out of our analyses had less aggressive tumors than those who were included, but were similar with respect to other possible risk factors. Genotype frequencies were also not significantly different between groups, suggesting that a substantial bias is unlikely to be present.

The median follow-up time in this study was 45 months (3.8 years), which may have been too short to observe associations with recurrence and survival if there is a true effect of the SNPs on these outcomes. We did not correct for multiple comparisons and have not performed a replication study. The types of breast cancer patients treated at our institution and community-based facilities could be different with respect to tumor characteristics, possibly leading to our study population having a greater proportion

of aggressive tumor characteristics than would be expected from the source population, which could affect the generalizability of our findings. It is possible, although less likely, that there could be differences with respect to genotype as well.

Conclusions

In conclusion, we showed that SNPs in *BRMS1* and *SIPA1* may be associated with tumor characteristics related to prognosis. Additional studies are needed to validate these findings and further investigate relationships between genetic variation in metastasis-modifying genes and the metastatic phenotype. Understanding the biology of metastasis and identifying biomarkers of recurrence are necessary to improve prediction of the subset of patients who will experience metastatic relapse, particularly since treatment for breast cancer often results in significant patient morbidity. While many women will never progress to metastatic disease, identifying factors associated with metastatic recurrence is critical to achieving effective, efficient therapy for breast cancer patients.

Acknowledgments This research was supported by a gift from the Jayne and Philip Hubbell Family. MRR is supported by DAMD W81XWH-11-1-0024, a predoctoral fellowship from the Department of Defense Breast Cancer Research Program, and previously supported by R25-CA11395102, an NCI predoctoral award. CBA is a recipient of funding from the Breast Cancer Research Foundation. Biospecimens and data were obtained from the RPCI DataBank and Biorepository, and genotyping was performed in the RPCI Genomics Core Facility, both of which are Cancer Center Support Grant Shared Resources supported by National Institutes of Health P30 CA016056-27. The funding agents played no role in any part of the study, including the design, collection, analysis and interpretation of the data, in the writing of the manuscript, or the decision to submit the manuscript for publication.

Conflict of interests The authors declare that they have no conflict of interests.

Ethical standard This work complies with all ethical standards and current laws of the USA.

References

1. Cancer facts and figures (2013) American Cancer Society
2. Higgins MJ, Wolff AC (2008) Therapeutic options in the management of metastatic breast cancer. Oncology 22:614–623 discussion 623, 627–9
3. Nguyen DX, Bos PD, Massagué J (2009) Metastasis: from dissemination to organ-specific colonization. Nat Rev Cancer 9:274–284
4. Nathanson SD, Kwon D, Kapke A, Alford SH, Chitale D (2009) The role of lymph node metastasis in the systemic dissemination of breast cancer. Ann Surg Oncol 16:3396–3405
5. Tresserra F, Rodriguez I, García-Yuste M, Grases PJ, Ara C, Fabregas R (2007) Tumor size and lymph node status in multifocal breast cancer. Breast J 13:68–71
6. Koscielny S, Arriagada R, Adolfsson J, Fornander T, Bergh J (2009) Impact of tumour size on axillary involvement and distant dissemination in breast cancer. Brit J Cancer 101:902–907
7. Payne SJL, Bowen RL, Jones JL, Wells Ca (2008) Predictive markers in breast cancer—the present. Histopathology 52:82–90
8. Stafford LJ, Vaidya KS, Welch DR (2008) Metastasis suppressors genes in cancer. Int J Biochem Cell B 40:874–891
9. Hedley BD, Vaidya KS, Phadke P, MacKenzie L, Dales DW, Postenka CO, MacDonald IC, Chambers AF (2008) BRMS1 suppresses breast cancer metastasis in multiple experimental models of metastasis by reducing solitary cell survival and inhibiting growth initiation. Clin Exp Metastas 25:727–740
10. Seraj MJ, Samant RS, Verderame MF, Welch DR (2000) Functional evidence for a novel human breast carcinoma metastasis suppressor, BRMS1, encoded at chromosome 11q13. Cancer Res 60:2764–2769
11. Phadke Pa, Vaidya KS, Nash KT, Hurst DR, Welch DR (2008) BRMS1 suppresses breast cancer experimental metastasis to multiple organs by inhibiting several steps of the metastatic process. Am J Path 172:809–817
12. Liu Y, Smith PW, Jones DR (2006) Breast cancer metastasis suppressor 1 functions as a corepressor by enhancing histone deacetylase 1-mediated deacetylation of RelA/p65 and promoting apoptosis. Mol Cell Biol 26:8683–8696
13. Hicks DG, Yoder BJ, Short S, Tarr S, Prescott N, Crowe JP, Dawson AE, Budd GT, Sizemore S, Cicek M, Choueiri TK, Tubbs RR, Gaile D, Nowak N, Accavitti-Loper MA, Frost AR, Welch DR, Casey G (2006) Loss of breast cancer metastasis suppressor 1 protein expression predicts reduced disease-free survival in subsets of breast cancer patients. Clin Cancer Res 12:6702–6708
14. Zhang Z, Yamashita H, Toyama T, Yamamoto Y, Kawasoe T, Iwase H (2006) Reduced expression of the breast cancer metastasis suppressor 1 mRNA is correlated with poor progress in breast cancer. Clin Cancer Res 12:6410–6414
15. Kelly LM, Buggy Y, Hill A, O'Donovan N, Duggan C, McDermott EW, O'Higgins NJ, Young L, Duffy MJ (2005) Expression of the breast cancer metastasis suppressor gene, BRMS1, in human breast carcinoma: lack of correlation with metastasis to axillary lymph nodes. Tumour Biol 26:213–216
16. Stark AM, Tongers K, Maass N, Mehdorn HM, Held-Feindt J (2005) Reduced metastasis-suppressor gene mRNA-expression in breast cancer brain metastases. J Cancer Res Clin 131:191–198
17. Lombardi G, Di Cristofano C, Capodanno A, Iorio MC, Aretini P, Isola P, Tancredi M, Collecchi P, Naccarato AG, Porta RP, Bevilacqua G, Caligo MA (2007) High level of messenger RNA for BRMS1 in primary breast carcinomas is associated with poor prognosis. Int J Cancer 120:1169–1178
18. Tsukamoto N, Hattori M, Yang H, Bos JL, Minato N (1999) Rap1 GTPase-activating protein SPA-1 negatively regulates cell adhesion. J Biol Chem 274:18463–18469
19. Crawford NPS, Walker RC, Lukes L, Officewala JS, Williams RW, Hunter KW (2008) The diasporin pathway: a tumor progression-related transcriptional network that predicts breast cancer survival. Clin Exp Metastas 25:357–369
20. Park YG, Zhao X, Lesueur F, Lowy DR, Lancaster M, Pharoah P, Qian X, Hunter KW, Yang H (2005) Sipa1 is a candidate for the metastasis efficiency modifier locus Mtes1. Nat Genet 37:1055–1062
21. Crawford NPS, Ziogas A, Peel DJ, Hess J, Anton-Culver H, Hunter KW (2006) Germline polymorphisms in SIPA1 are associated with metastasis and other indicators of poor prognosis in breast cancer. Breast Cancer Res 8:R16

22. Gaudet MM, Hunter K, Pharoah P, Dunning AM, Driver K, Sherman M, Peplonska B, Brinton LA, Lissowska J, Chanock S, Garcia-Closas M (2009) Genetic variation in SIPA1 in relation to breast cancer risk and survival after breast cancer diagnosis. Int J Cancer 124:1716–1720
23. Hsieh S-M, Look MP, Sieuwerts AM, Foekens Ja, Hunter KW (2009) Distinct inherited metastasis susceptibility exists for different breast cancer subtypes: a prognosis study. Breast Cancer Res 11:R75
24. Ambrosone CB, Nesline MK, Davis W (2006) Establishing a cancer center data bank and biorepository for multidisciplinary research. Cancer Epidem Biomar 15:1575–1577
25. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–311
26. Meehan WJ, Welch DR (2003) Breast cancer metastasis suppressor 1: update. Clin Exp Metastas 20:45–50
27. Meehan WJ, Samant RS, Hopper JE, Carrozza MJ, Shevde LA, Workman JL, Eckert KA, Verderame MF, Welch DR (2004) Breast cancer metastasis suppressor 1 (BRMS1) forms complexes with retinoblastoma-binding protein 1 (RBP1) and the mSin3 histone deacetylase complex and represses transcription. J Biol Chem 279:1562–1569